

## BAB VI

### PENUTUP

#### 6.1 Kesimpulan

Berdasarkan hasil penelitian, implementasi, dan analisis perbandingan dalam klasifikasi penyakit *Tuberkulosis*, maka dapat ditarik kesimpulan sebagai berikut:

1. Algoritma *Random Forest* secara konsisten mengungguli *K-Nearest Neighbors* (KNN) pada seluruh skenario pengujian. Pada konfigurasi terbaiknya (15 Fitur, 5-Fold), *Random Forest* mencapai tingkat *Accuracy* sebesar 99,0093%, sedangkan KNN pada skenario yang sama hanya mencapai 98,2593%. Hal ini menunjukkan bahwa metode *ensemble* pada *Random Forest* lebih efektif dalam mengolah pola data klinis *Tuberkulosis* dibandingkan pendekatan berbasis jarak pada KNN.
2. Penambahan jumlah fitur dari 10 menjadi 15 fitur terbukti meningkatkan performa model klasifikasi. Berdasarkan hasil pengujian, penggunaan 10 fitur menghasilkan akurasi tertinggi sebesar 98,5123% pada algoritma *Random Forest* dengan *5-Fold Cross Validation*, sedangkan KNN pada konfigurasi yang sama mencapai akurasi 98,0123%. Sementara itu, penggunaan 15 fitur memberikan peningkatan performa yang lebih baik, dengan akurasi tertinggi mencapai 99,0093% pada algoritma *Random Forest* dengan *5-Fold Cross Validation*, dan 98,2593% pada KNN.
3. Implementasi *5-Fold Cross Validation* menunjukkan hasil yang lebih optimal bagi algoritma *Random Forest* dengan 15 fitur dibandingkan *10-Fold*. Penggunaan *5-Fold* memberikan keseimbangan yang baik dalam pembagian data latih dan data validasi, sehingga model mencapai akurasi puncak (99,0093%). Sebaliknya, pada *10-Fold*, meskipun hasilnya masih sangat baik (98,7744%), terdapat sedikit penurunan yang mengindikasikan bahwa pembagian data ke subset yang terlalu kecil cenderung meningkatkan sensitivitas varians pada model.
4. Penerapan teknik *Majority Voting* menghasilkan akurasi akhir gabungan sebesar 98,00%, dengan 48 dari 50 data uji berhasil diklasifikasikan dengan benar. Hasil ini menunjukkan bahwa model mampu mengoreksi variasi

prediksi antar-*fold* secara efektif serta menghasilkan prediksi yang konsisten. Tingginya tingkat akurasi dan konsistensi tersebut membuktikan bahwa algoritma *Random Forest* memiliki reliabilitas yang sangat baik dan layak digunakan sebagai sistem pendukung keputusan medis dalam klasifikasi status *Tuberkulosis* pada data riil di lapangan.

## 6.2 Saran

Untuk pengembangan penelitian selanjutnya, disarankan beberapa hal sebagai berikut:

1. Perluasan dan penyeimbangan *dataset* penyakit paru mengingat keterbatasan jumlah *dataset* saat ini yang didominasi oleh kasus *pneumonia* untuk kategori non-TBC, disarankan agar penelitian berikutnya menambah volume data dengan variasi penyakit paru lainnya, seperti *Bronkitis*, Asma, atau PPOK. Hal ini sangat penting untuk mengatasi masalah ketimpangan data (*imbalanced data*) dan melatih model agar lebih sensitif dalam membedakan gejala *tuberkulosis* dengan penyakit paru lain yang memiliki kemiripan klinis.
2. Penambahan atribut dan gejala spesifik disarankan untuk menambahkan fitur atau atribut gejala yang lebih mendalam, seperti riwayat kontak erat dengan penderita TBC, kondisi lingkungan tempat tinggal, atau hasil pemeriksaan penunjang medis lainnya. Penambahan variabel ini bertujuan untuk meningkatkan kemampuan diskriminasi model dalam menangani kasus-kasus diagnosa yang lebih kompleks.
3. Pengembangan Platform Aplikasi Praktis Hasil klasifikasi ini diharapkan dapat dikembangkan ke dalam bentuk aplikasi berbasis mobile atau web. Dengan adanya platform tersebut, sistem ini dapat digunakan secara nyata oleh masyarakat sebagai sarana deteksi dini mandiri, maupun oleh tenaga kesehatan di Puskesmas untuk membantu mempercepat proses skrining awal pasien.